3-1986

Is Statistical Discrimination Efficient?

Stewart J. Schwab
Cornell Law School, sjs15@cornell.edu

Is Statistical Discrimination Efficient?

By STEWART SCHWAB*

Neoclassical economists have advanced two general types of labor market discrimination models:[1] taste discrimination models and statistical discrimination models. Taste models include in the utility functions of employers, fellow workers or customers a desire to avoid members of certain groups. (See Gary Becker, 1957.) Under such an approach, discrimination cannot be characterized as either efficient or inefficient. One can compare the distribution of income and utility with an economy where people do not have a taste for discrimination. But until one decides the moral question of whether furthering the taste is acceptable, one cannot begin to ask whether society's resources are being placed in their most productive uses.

Statistical discrimination, by contrast, can affect efficiency. Statistical discrimination differs from the classic taste-for-discrimination model in assuming no prejudice or invidious motive by employers or employees, but rather that employers use average characteristics of groups to predict individual worker attributes.[2] The early models examined only the distributive consequences of statistical discrimination (see Kenneth Arrow, 1972, 1973; Edmund Phelps, 1972; Dennis Aigner and Glen Cain, 1977; and George Borjas and Matthew Goldberg, 1978). The literature tacitly assumed that statistical discrimination, whereby firms use valid (and free) information solely to maximize profits, must be more efficient than an economy where firms ignore the information. The standard statistical-discrimination model thus presents society with an uncomfortable tradeoff. In prohibiting statistical discrimination, society must accept lower national output.

Recently, Shelly Lundberg and Richard Startz (1983) expanded the standard model of Aigner-Cain to examine some of the efficiency effects of statistical discrimination. Aigner and Cain had considered a labor market where employers can only imperfectly test worker ability, and the test predicts more accurately for workers of one group than for another. Aigner and Cain had derived from this model an employer wage scale that offers a worker a convex combination of the mean productivity of his group and his individual test score, and had examined the distributional consequences between groups of such an offer. Lundberg and Startz examined how these wage offers might affect human capital decisions. They first showed that workers in such a world invest too little in training, because individual workers are not completely compensated for individual increases in productivity. Lundberg and Startz then concluded that statistical discrimination may exacerbate the inefficiencies. Although the wage of the favored group rises closer to the allocatively efficient wage, the wage of the disfavored group falls further from the efficient wage.

*Cornell Law School, Ithaca, NY 14853. This paper is a substantially revised section of my doctoral dissertation. I thank Peter Steiner and my dissertation committee members Frank Stafford, Glenn Loury, Theodore St. Antoine, and Gavin Wright for their guidance. I also acknowledge comments by Ron Ehrenberg, Olivia Mitchell, and other participants at the Labor Economics Workshop of Cornell's New York State School of Industrial and Labor Relations, and an anonymous referee.

[1]A third basic framework, somewhat outside the neoclassical tradition, uses labor market segmentation or "dual" labor markets to explain the occurrence and persistence of discrimination.

[2]I sometimes refer to statistical discrimination as a situation where an employer acts on a "true stereotype." Examples of stereotypes include "blacks are less skilled than whites," "women quit more frequently than men," and "women live longer than men." The term stereotype reflects society's moral distaste for statistical discrimination, under the battle cry "judge me, not my group." Yet the word "true" is a crucial modifier in the phrase true stereotype. The employer, I assume, responds only to correct group information (statements that are indeed true on average) because competitive forces will eliminate employer decisions based on false information. See my dissertation for an extended legal and moral evaluation of statistical discrimination.

In the present paper, I likewise examine the efficiency effects of statistical discrimination. I start from a different strain in the imperfect-information literature, the George Akerlof (1970) and Hayne Leland (1979) "lemons" model. Adapting this model to the labor market, I examine whether employers' use of group information will decrease allocative inefficiencies in labor supply. As Lundberg and Startz found for human capital investments, I find that statistical discrimination increases efficiency of labor supply for the favored group but decreases efficiency for the disfavored group. My general model outlines the parameters and suggests that the net efficiency effect cannot be determined a priori. Importantly, my model suggests that statistical discrimination *can* reduce the efficiency of the economy even if the two groups differ in their underlying productivities. In Section III, a variant of the model more strongly concludes that statistical discrimination *will* exacerbate inefficiencies under certain conditions.

I. Labor Supply Distortions from Limited Information

Statistical discrimination would not occur if employers knew the ability of individual workers. In many settings, however, employers cannot obtain individual information, so they resort to less precise group information to fill the information void. To determine whether "filling the void" increases allocative efficiency, I first examine the efficiency losses from imperfect information. In Section II, I give employers group information to fill the void.

A. *Labor Demand*

Employers cannot distinguish among individuals, so they set the wage rate at the average product of employed workers. For simplicity, I assume that output equals the amount of effective labor used, uF, where u is the average ability of employed workers and F is the number of employed workers. In a competitive market, then, the demand relation for labor quality is

$$u = w. \tag{1}$$

B. *Labor Supply*

Individuals can work in one of two markets. In the "standardized" job market described above, employers cannot distinguish among workers. All workers receive the same wage. The alternate market—the "individualized" market—can identify and pay workers according to individual productivity.[3] This market includes self-employed persons (including homemakers), persons paid by piece work, entrepreneurs, and managers for whom the bottom line reflects individual managerial ability. The key distinction is that employers in the standardized market have an incentive to discriminate statistically, whereas employers in the individualized market know individual ability and thus need not rely on less precise group information.[4]

A critical assumption is that the more skilled workers in the standardized market can also produce more in the individualized market.[5] As the standardized wage increases, more-able workers are drawn into this market. The last worker necessarily has a higher marginal product than the average worker (and workers in the individualized market still higher marginal products).

To model this more formally, let a be an index over the interval $(0,Z)$ of an individual's productivity in the standardized market. Let $f(a)$ be the number of workers of ability a. A worker supplies one unit if he decides to work in the standardized market. A person will work if the standardized-market wage, w, exceeds his individualized-market productivity, P. To capture the idea that more-able standardized workers can also produce more in the individualized market, I assume the

[3] To avoid general equilibrium problems of shifting prices, I assume that workers produce the same product in the standardized and individualized market and that wages are paid in output product units.

[4] See Michael Spence (1974a, b) for a discussion of distinctions between markets where employers can observe individual productivity and markets where signalling and statistical discrimination will occur.

[5] The assumption is equivalent to Akerlof's and Leland's assumption that as the market price of goods rises, higher quality goods are sold. See also James Heckman (1974), who deals with the wage/quality issue in the labor market.

following relation:

(2) $$P = P(a) \quad P'(a) > 0, \quad P''(a) > 0.$$

The positive second derivative suggests that, at some point, individual productivity in a collective setting has some limits.

Define A as the highest ability level appearing in the standardized market. Given the labor supply assumption, we have $A = A(w)$, where $A(w)$ is the inverse function of P. The number of workers willing to work in the standardized market for a given wage is

(3) $$F(w) = \int_0^{A(w)} f(a)\, da.$$

Average ability in the standardized market, u, increases with w:

(4) $$u(w) = \int_0^{A(w)} af(a)\, da / F(w).$$

Note that the supply equations for the quantity and quality of labor (equations (3) and (4)) are dependent on each other.

C. *Suboptimal Equilibrium*

Substituting equations (3) and (4) into equation (1) creates a single equilibrium equation in w:[6]

(5) $$u(w) = w.$$

Figure 1 graphs the equilibrium wage and

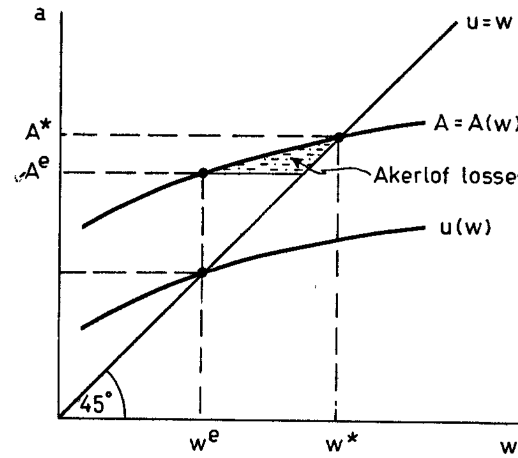

FIGURE 1. EFFICIENCY LOSSES WITH NO INFORMATION ABOUT INDIVIDUAL WORKER ABILITY

average ability, w^e and u^e. Overall social product is the sum of production in the standardized and individualized market (call this *GNP*):

(6) $$GNP = \int_0^{A(w)} af(a)\, da + \int_{A(w)}^{Z} P(a) f(a)\, da.$$

GNP would be maximized by allocating workers to equate the marginal output of standardized and individualized market workers (point A^*, w^* in Figure 1), with all workers of ability less than A^* working in the standardized market. As can be seen from Figure 1, an unregulated economy allocates too few workers to the standardized market.[7] The problem is that firms value the (above average) marginal worker as a person of average ability, because firms must treat all persons on average. This discourages able

[6]Stability requires that $u(w)$ be concave in the relevant range, a condition I assume throughout the paper. Differentiating equation (4), and dropping the arguments of the functions for brevity, yields $u'(w) = fA'(A-u)/F > 0$. Differentiating again yields $u''(w) = fA''(A-u)/F + [A']^2 f'(A-u)/F + fA'(A'-u)/F + [fA']^2(u-A)/F^2$. The last term is always negative. Term 1 is negative since, by the second-order restrictions of equation (2), $A'' < 0$. Term 2 is negative if $f' < 0$, which occurs for all unimodal distributions whenever A exceeds the mode. Term 3 is negative if $f(A-u) > F$, which may well not be true. Although there is some indeterminateness for the sign of u'' in general, u is concave for a wide range of ability distributions. This holds, for example, for all ability distributions of the form $f(a) = ca^n$, $n > -1$ or < -2, since $u'' = A''(n+1)/(n+2)$.

[7]The government could achieve first-best social efficiency by subsidizing the wages of all workers, with an optimal subsidy being $S^* = A^* - u^*$. This possibility is analogous to the government policy mentioned by Lundberg and Startz (p. 344, fn. 6) of subsidizing wages to induce optimal human capital investment by workers. As they point out, in practice it would be extremely difficult to calculate and implement this subsidy.

persons from working in the standardized market. Indeed, in equilibrium, the marginal worker can produce in the individualized market what the average person produces in the standardized market. The problem of limited information is thus the gap between the marginal contributions of the average and marginal worker.[8] Following the original lemons model, I call the losses caused by this gap "Akerlof losses." (See Figure 1.)

II. Adding Statistical Discrimination to the Model

Suppose now that firms in the standardized market know something about individual workers. Our question is whether this information will reduce the Akerlof inefficiencies. To create a statistical-discrimination model, I give firms additional information in the form of a true stereotype.

Workers belong to one of two groups, which firms can costlessly identify. Within each group, workers vary in ability. All workers of ability a have the same labor market behavior, regardless of the group to which they belong. For reasons exogenous to the model, however, workers of group 1 have a higher average ability in the standardized market than workers of group 2, for any given wage. The true stereotype, then, is that "group 1 workers are more productive than group 2 workers."

As before, production in the standardized market equals the amount of effective labor, $L = u^1F^1 + u^2F^2$. Firms using the group information will treat all persons within a group equally, and will offer a wage of $w^1 = u^1$ to group 1 persons, and a lower wage of $w^2 = u^2$ to group 2 persons. The quantity and quality of labor supplied for each group are analogous to equations (3) and (4). Figure 2 graphs the separating equilibrium.

Our question is whether statistical discrimination increases output. If firms use group information, GNP^d (GNP discriminatory) is

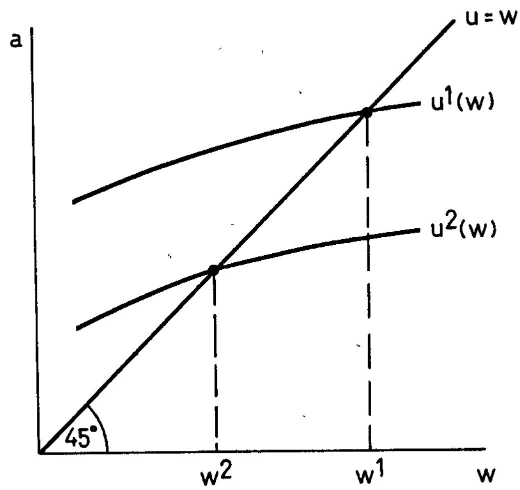

FIGURE 2. EQUILIBRIUM WAGES WITH GROUP INFORMATION

$$GNP^d = \int_0^{A^1 = A(w^1)} af^1(a)\,da + \int_0^{A^2 = A(w^2)} af^2(a)\,da + \int_{A^1}^{Z} P(a)f^1(a)\,da + \int_{A^2}^{Z} P(a)f^2(a)\,da. \tag{7}$$

Subtracting (7) from (6), statistical discrimination increases social output if and only if

$$\int_{A}^{A^1}[a - P(a)]\,f^1(a)\,da > \int_{A^2}^{A}[a - P(a)]\,f^2(a)\,da. \tag{8}$$

As equation (8) shows, the efficiency of statistical discrimination depends on three factors. First is the relative number of group 1 and group 2 persons who are highly skilled, marginal workers. Second is the extent of the shift in wages. Third is the net change in social output as group 1 workers enter the standardized market and group 2 workers leave for the individualized market. The issue is whether the gap in the marginal worker's production in the two markets rises or falls with ability over the relevant range. In terms

[8] In a similar vein, Spence (1975) has analyzed the problem of the gap between the valuation of quality by the average and marginal consumer.

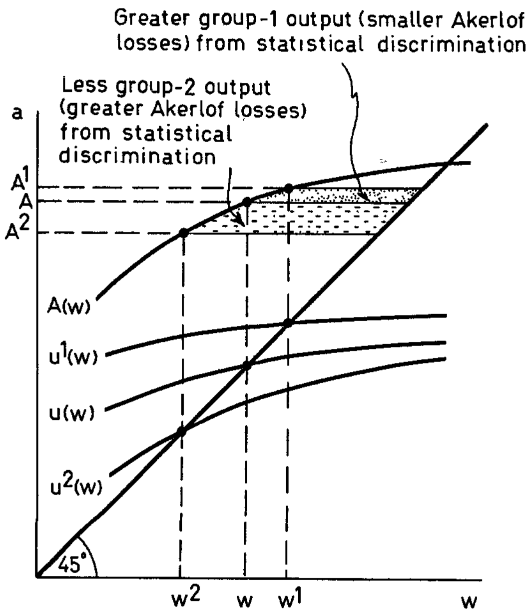

FIGURE 3. EFFICIENCY COMPARISON OF DISCRIMINATORY AND NONDISCRIMINATORY EQUILIBRIA

of the model, this third factor will tend to make statistical discrimination efficient if $P' > 1$.

Figure 3 sketches the gains and losses in efficiency from the second and third factors. The $A(w)$ schedule, showing the ability of the most-able standardized-market worker for a given wage, is identical for both groups since workers of the same ability behave identically regardless of group. It is along this $A(w)$ curve of marginal workers that the relative efficiency calculations are made. But Figure 3 also shows that the wages of the two groups, which determine the extent of the relative Akerlof losses, are based on non-marginal group averages. The wages thus have no necessary relation to efficiency. A ban on statistical discrimination pushes the wage of group 2 workers closer to a socially optimal incentive to work, but causes the wage of group 1 workers to diverge further from an efficient level. The net efficiency effect depends on the shape of the ability functions of the two groups. As drawn, Figure 3 suggests that statistical discrimination decreases output. The large loss in group 2 average ability outweighs the increase in group 1 average ability. Note that the only restriction on the underlying ability distributions is that u^1 exceed u^2 over the relevant range. Thus, for example, statistical discrimination can be inefficient even if the groups differ in the underlying average ability.

Statistical discrimination is most likely to be inefficient when the disfavored group has relatively large numbers of unskilled workers, holding down the average ability of the group, while the skilled workers are more evenly dispersed between groups. Figure 3 was drawn with this distribution in mind: the average abilities of the two groups at low levels of the truncated distribution (u^1 and u^2 at low wages) vary widely, but the gap in average ability narrows as more-skilled workers enter the labor market.

III. The Inefficiency of Statistical Discrimination when Labor Supply Elasticity Differs by Group

The previous model treated the standardized/individualized market choice of workers as invariant across groups. The group differences that employers observed arose from differences in ability distributions ($f^1 \neq f^2$). Group differences can also occur when the opposite assumptions are made: standardized ability distributions are the same for groups 1 and 2 but the reservation-wage functions (indicating individualized-market productivity) differ by group.

Suppose the standardized labor supply of group 1 is highly inelastic, so that virtually all members of group 1 work in the standardized market for any wage in the relevant range. In other words, A^1 and u^1 are nearly constant functions of the wage. Suppose the supply elasticity of group 2 is greater: indeed, assume that the supply response of group 2 persons is analogous to equations (3) and (4) of the previous model. In this situation, employers in the standardized market will observe that the average productivity of group 1 exceeds that of group 2, because virtually all group 1 persons appear in the standardized market but the most-able group 2 persons remain in the individualized market.

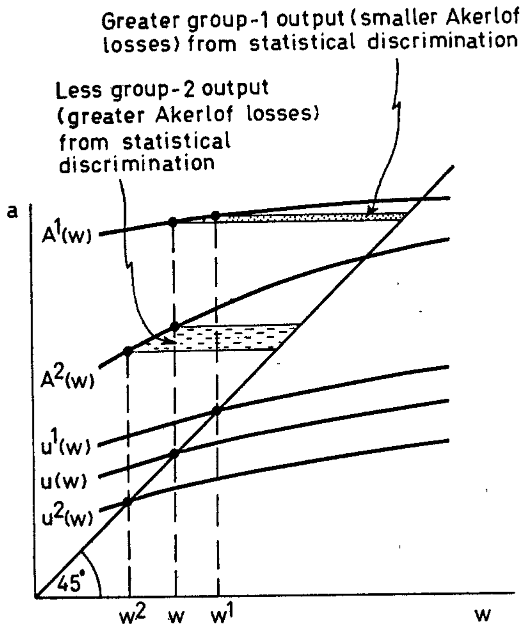

FIGURE 4. INEFFICIENCY WHEN LABOR SUPPLY ELASTICITY DIFFERS BY GROUP

Figure 4 graphs this variant of the model. The important result is that statistical discrimination is now *unambiguously inefficient.* Most group 1 workers will work regardless of the wage, so increasing their wage from w to w^1 does not induce a larger supply of workers into the standardized market. Lowering the group 2 wage from w to w^2, however, discourages talented persons from working in the standardized market, and thus lowers output. Society is more productive in this situation if firms ignore group information.

IV. Concluding Remarks

I have answered the question of the title of this piece with a definite "maybe not." Even though statistical discrimination is based on free, accurate information, it does not necessarily allocate resources more efficiently than if firms ignored the information. I have identified two situations in which firms' use of group information may exacerbate the labor supply distortions of limited information. First, one group may have many unskilled persons who lower the expected ability of all persons in the group. The skilled workers at the margin, however, may be more evenly dispersed between groups. If so, using group information to set wages will discourage more able workers than it encourages. Societal output decreases. Second, labor supply inefficiencies arise if the reservation-wage function differs between groups. In such a case, wages set by group information will discourage group 2 persons without an offsetting encouragement of group 1 persons.

To be certain of the inefficiency of statistical discrimination in specific situations, one needs empirical documentation. Unfortunately, empirical testing of statistical discrimination (and, indeed, of the signalling hypothesis in general) is difficult.[9] Further refinement of the models should enable empirical testing. In the meantime, the negative theorem remains important in shifting the burden of proof on the acceptability of statistical discrimination. Firms commonly assert that they must statistically discriminate to maximize profits, which—as Adam Smith showed—benefits society in general. The negative theorem indicates, however, that an a priori efficiency claim cannot be used to justify statistical discrimination. Statistical discrimination may indeed be socially inefficient. Perhaps we should thus be more confident in acting upon our moral approbation of it, at least until empirical documentation demonstrates in which direction the efficiencies lie.

[9] For one effort, see Daniel Dick and Marshall Medoff (1976).

REFERENCES

Aigner, Dennis and Cain, Glen, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, January 1977, *30*, 175–87.

Akerlof, George, "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, August 1970, *84*, 488–500.

Arrow, Kenneth, "Some Mathematical Models of Race Discrimination in the Labor

Market," in Anthony Pascal, ed., *Racial Discrimination in Economic Life*, Lexington: Lexington Books, 1972.

______, "The Theory of Discrimination," in Orley Ashenfelter and Albert Rees, eds., *Discrimination in Labor Markets*, Princeton: Princeton University Press, 1973.

Becker, Gary, *The Economics of Discrimination*, Chicago: University of Chicago Press, 1957.

Borjas, George and Goldberg, Matthew, "Biased Screening and Discrimination in the Labor Market," *American Economic Review*, December 1978, *68*, 918–22.

Dick, Daniel and Medoff, Marshall, "Filtering by Race and Education in the U.S. Manufacturing Sector: Constant-Ratio Elasticity of Substitution Evidence," *Review of Economics and Statistics*, May 1976, *58*, 148–55.

Heckman, James, "Shadow Wages, Market Wages, and Labor Supply," *Econometrica*, July 1974, *42*, 679–94.

Leland, Hayne, "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards," *Journal of Political Economy*, December 1979, *87*, 1328–46.

Lundberg, Shelly and Startz, Richard, "Private Discrimination and Social Intervention in Competitive Labor Markets," *American Economic Review*, June 1983, *73*, 340–47.

Phelps, Edmund, "The Statistical Theory of Racism and Sexism," *American Economic Review*, September 1972, *62*, 659–61.

Schwab, Stewart, "Stereotypes, Imperfect-Information Theories, and Statistical Discrimination in Labor Markets," unpublished doctoral dissertation, University of Michigan, 1981.

Spence, A. Michael, (1974a) *Market Signalling: Information Transfer in Hiring and Related Screening Processes*, Cambridge: Harvard University Press, 1974.

______, (1974b) "Competitive and Optimal Responses to Signals: An Analysis of Efficiency and Distribution," *Journal of Economic Theory*, March 1974, *7*, 296–332.

______, "Monopoly, Quality, and Regulation," *Bell Journal of Economics*, Autumn 1975, *6*, 417–29.